(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification[7]: G01N 33/68

(21) International Application Number: PCT/GB03/00112

(22) International Filing Date: 13 January 2003 (13.01.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0200595.7    11 January 2002 (11.01.2002)   GB

(71) Applicant (for all designated States except US): OXFORD GLYCOSCIENCES (UK) LTD [GB/GB]; Attn: Mary Gadsden, The Forum, 86 Milton Park, Abingdon, Oxfordshire OX14 4RY (GB).

(72) Inventor; and
(75) Inventor/Applicant (for US only): SOLOVIEV, Mikhail [GB/GB]; The Forum, 86 Milton Park, Abingdon, Oxfordshire OX14 4RY (GB).

(74) Agent: BLAKEY, Alison, Jane; Oxford GlycoSciences (UK) Ltd, The Forum, 86 Milton Park, Abingdon, Oxfordshire OX14 4RY (GB).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— with international search report
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: METHOD FOR QUANTITATION OF PROTEIN LEVELS IN BIOLOGICAL SAMPLES

(57) Abstract: The invention provides a novel method for the quantitation of differential levels of proteins using unlabelled samples in particular using protein binding arrays. The method is useful for assessing differential expression or post-translational modification of proteins, and in diagnosis and screening where differential expression or post-translational modification of proteins is an indicator of the diseased state.

## Method for Quantitation of Protein Levels In Biological Samples

The invention disclosed provides a novel method for the quantitation of differential levels of proteins using unlabelled samples in particular using protein binding arrays. The
5  method is useful for assessing differential expression or post translational modification of proteins, and in diagnosis and screening where differential expression or post translational modification of proteins is an indicator of the diseased state. The method can be extended to permit the use of generic arrays.

Characterization of the complement of expressed proteins from a single genome is a
10  central focus of the evolving field of proteomics. Since one genome produces many proteomes (hundreds in multi-cellular organisms) and the number of expressed genes in a cell is minimally 10,000, the characterization of thousands of proteins to evaluate proteomes can only be readily accomplished using a high-throughput, automated process.

Monitoring the expression of, and properties of a large number of proteins provides
15  important information about the physiological or biochemical state of a cell. A cell can express a large number of different proteins, and the expression patterns (the number of proteins expressed and the expression levels) vary in different cell types, explaining why different cells perform different functions. Since many diseases associate with, or even result from a change in the protein expression pattern, comparing protein expression
20  patterns between normal and disease conditions can reveal proteins whose changes are critical in the disease and thus can identify proteins which are of therapeutic value or which aid diagnosis. Methods of detecting protein expression profiles also have important applications in, for example and without limitation, tissue typing, forensic identification, toxicology testing and clinical diagnosis. Detecting patterns of post translational modification
25  can convey similarly useful information.

Microarray formats for the quantitative detection of proteins (including antibody and antigen capture) have been developed for diagnosis (Ekins & Chu, 1999, Trends in Biotechnol. 17:217-8) and protein-protein interaction discovery (Walter et al., 2000, Curr. Opin. Microbiol. 3:298-302; and US 6,197,599). However, development of a microarray
30  affinity capture system for semi-quantitative or quantitative proteomics similar to those used for semi-quantitative mRNA expression analyses (Schena et al., 1998, Trends in Biotechnol. 16:301-6) has lagged behind.

Although mRNA microarrays are well advanced and identification of the differential sequences is straightforward, there is a large inconsistency between mRNA and protein
35  levels (Gygi, SP et al., 1999, Mol. Cell Biol. 19:1720-30) and since proteins are the primary agents responsible for disease and drug responses, predicting protein presence and amount by measurement of mRNA can lead to significant errors, sometimes as much as 20-fold or larger. Even more important is that activities of many proteins can be grossly affected by post-translational modifications such as phosphorylation. A nucleic acid array cannot detect
40  such effects. Also current proteomics techniques such as 2D-gel electrophoresis require complex procedures to identify differentially expressed proteins (Parekh et al. 2000, Journal of Commercial Biotechnology 6:284-291). Most significantly, low abundance proteins can escape detection due to limitations in visualizing the proteins present in the 2D-gel. A microarray format, high throughput system for quantitative direct detection of proteins such

as is required in proteomics applications is thus highly desirable for, without limitation, diagnosis, pharmacoproteomics, identification of markers of disease and drug target discovery.

5      Current array-based proteomics techniques are in their infancy and suffer from a variety of problems including the isolation of sufficient numbers of high affinity capture agents cognate (i.e. recognising) for the proteins of interest. A more pressing problem however, is the detection of an interaction and, particularly, detecting differences in protein expression between two or more samples. Detection of cDNAs in an expression analysis using DNA microarrays is via fluorescent labelling during the reverse transcription step. No 10    such step exists is the use of protein binding arrays however, as generally fluorescent labelling of proteins using chemical means is used.

Protein binding arrays known in the art suffer from multiple drawbacks (see Kodadek, T., 2001, Chemistry & Biology 8:105-115 for a discussion). In particular, there are non-specific binding and background issues where high levels of label e.g. fluorescent label, can 15    obscure specific binding. There are also sample labelling issues where the inherent heterogeneity of proteins, polypeptides and peptides means that they can all take up label to different extents making comparison of expression levels impossible (Abbot, A., 2002, Nature 415:112-114). Inherent in labelling a sample is the need to purify the sample free from unbound reactive label. This consequently results in precious sample losses during 20    purification and necessitates relatively large starting sample requirements, a problem that researchers have not been able to overcome even using cDNA microarrays (Kling, J., 2002, The Scientist 16: 51). Another principal difficulty highlighting the difference between protein-antibody interactions and nucleic acid complementary chain interactions is that protein or antibody labelling methods use reactive amino acid side chains to attach the label, for 25    example, a fluorescent dye. Labelling of such reactive side chains can compromise the binding of a protein to its cognate capture agent or interacting partner. In the case of antibodies, the labelling may occur within the variable domains and compromise antigenic recognition by said antibody. Thus, protein labelling results in chemical modification of the surface of the proteins and may decrease the affinity of an antibody-protein interaction or a 30    protein-protein or protein-nucleic acid interaction. Where the biomolecules of interest are small polypeptides or short peptides the introduction of a chemical modification in the form of a label may interfere such that total disruption of the binding of said biomolecules to cognate capture agents or other binding partners occurs. This is unlike the labelling of nucleic acids, where a label is usually introduced via an enzymatic reaction into a defined position (such as 35    the 3'- or 5'-end of a nucleic acid) and hence do not disrupt the interaction between complementary chains.

Another major difference between nucleic acid arrays and protein arrays is that for any arrayed nucleic acid, the affinity of its interaction with a corresponding complementary strand is a product of length up to a plateau value. The affinity of interaction is thus similar 40    for all nucleic acid fragments of similar lengths. In contrast, antibodies are inherently heterogeneous and hence their affinities for their respective antigens are often very different irrespective of size. In the context of an antibody array, these differences in affinities can result in a wide range of signal intensities even if equimolar amounts of a labelled sample

are added to an array of cognate antibodies. Moreover, highly abundant proteins may fully saturate the immobilized antibody and mask any differential binding signal between two samples. The latter problems taken together, mean that any apparent differences in the binding signals obtained between two different samples, for example diseased versus

5    normal tissue, cannot be used to indicate the respective abundances of the proteins in the two samples as the relationship between the binding signal strength and the protein concentration is meaningless in the absence of a standard or a calibration curve. The problem is exacerbated if the aim is to compare differential expression between different cell or tissue types where the scale involved negates the idea of generating calibration curves for

10   every protein-capture agent complex. Arraying different amounts of antibody in inverse proportions to the affinity of a given antibody is also impractical.

Competitive displacement assays are known in the art, for example, radioimmunoassay (RIA). RIA utilises the competitive binding between pure, labelled antigen and pure unlabelled antigen for an antibody which specifically recognises said

15   antigens to determine the amount of antigen in a sample (Wilson & Goulding, Principles and Techniques of Practical Biochemistry, 3$^{rd}$ Ed. 1986, Cambridge University press, p138-141). The latter method requires the use of a purified antigen labelled with minimal alteration of the immunoreactivity of said antigen in conjunction with the use of known amounts of unlabelled antigen in order that a calibration curve may be constructed. The calibration

20   curve is then used to measure the amount of antigen in samples treated similarly. A competitive displacement assay is described in US 4,895,809, but this assay does not permit the measurement of differential expression of one or more proteins of interest within two unlabelled samples, and without the use of a calibration curve.

Disclosed herein is a novel method applicable to arrayed antibodies, other capture

25   agents or interacting biomolecules irrespective in each case of their relative affinities and that permits the reliable assessment of differential protein expression in a complex sample even between different cell types or tissue types without the use of a calibration curve. The method is tolerant towards high levels of non-specific binding and does not require any potentially interaction-disrupting labelling of a test sample, such as but not limited to, a

30   reference or control sample, or a test sample. Most particularly, it enables the comparison of two or more unlabelled samples over a wide concentration range that has hitherto been impractical.

Put most generally, there is provided is a method of comparing the level of at least one biomolecule within a sample with the level of said biomolecule in at least one other

35   sample, the method comprising:

    (a)    contacting a substrate with a labelled reference sample in the presence of an unlabelled reference sample wherein said substrate comprises at least one immobilised protein binding unit;

    (b)    contacting at least one other replicate substrate with an identical labelled

40   reference sample of part (a) in the presence of an unlabelled test sample;

    (c)    detecting and comparing the displacement of labelled reference sample from the substrates of parts (a) and (b).

More particularly, the invention provides a method of comparing the level of a target protein within a first sample with the level of said target protein within a second or subsequent sample, the method comprising:

5       (a)     contacting a labelled reference sample containing said target protein in the presence of the first sample with an immobilised protein binding unit capable of capturing said target protein;

      (b)     for each second or subsequent sample contacting a labelled reference sample which is a replicate of that of step (a) in the presence of the second or

10              subsequent sample with a immobilised protein binding unit which is a replicate of that of step (a); and

      (c)     detecting the level of label present on the protein binding units employed in steps (a) and (b) and from a comparison of said levels, determining a comparison of said target protein levels within the first sample and within the

15              second or subsequent sample.

The method of the invention employs at least one protein binding unit (PBU) immobilised on any substrate known in the art, such as but not limited to, beads, derivatized glass slides or Hydrogel™ pads. The substrate may comprise a bead, a substantially planar

20 surface or any suitable substrate known in the art such as, but not limited to, metal, glass, plastic, Hydrogels™ (Packard BioSciences, Meridian, CT) or membranes, or those documented by Walter *et al.* in 2000 (Curr. Opin. Microbiol. 3: 298-302). A PBU refers to at least one capture agent. Where a plurality of PBUs are used they are preferably arranged in a specific, co-ordinated pattern e.g. in a grid format, such that the location of each binding

25 unit is known. Such a format is referred to herein as an affinity array.

PBUs may be minimal recognition units comprising capture agents which include antibodies, antibody fragments, proteins, polypeptides, peptides, receptors, other biomolecules or complementarity determining regions (CDRs) and even nucleic acids. Capture agents for the production of affinity arrays of PBUs include, but are not limited to,

30 antibodies such as, but not limited to, polyclonal, monoclonal, bispecific, humanized or chimeric antibodies, single chain antibodies, Fab fragments and F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic antibodies, and epitope-binding fragments of any of the above produced by means well-known in the art, as well as but not limited to, proteins, polypeptides, peptides, other biomolecules, aptamers or

35 complementarity determining regions (CDRs). In native IgG antibodies the variable heavy and light chains each contribute three regions called CDRs that are responsible for binding antigen. Each region is typically 7-20 amino acids long, e.g. 15-17 amino acids, and its sequence defines the specificity and affinity of that CDR for the antigen.

There is increasing evidence that these regions of CDR-peptides are capable of

40 autonomous specific binding to antigens, see for example Steinbergs, J. *et al.*, 1996, Hum. Antibodies Hybridomas 7: 106-112; William, W. *et al.*, 1991, J. Biol. Chem., 266: 5182-5190; Saragovi, H. *et al.*, 1991, Science 253: 792-795; Welling, W. *et al.*, 1991, J Chromatogr. 548: 235-242; and Levi, M. *et al.*, 1993, Proc. Natl. Acad. Sci. USA 90: 4374-4378; and hence they have been proposed as new generation antibody fragments with reduced

immunogenicity and as anti-viral molecules (see Sivolapenko, G. *et al.,* 1995, Lancet 346: 1662-1666; and Rossenu, S. *et al.,* 1997, J. Prot. Chem., 16, 499-503). CDRs can be produced by recombinant means or can be chemically synthesized. CDRs can be chemically synthesized according to known CDR sequences using either standard protein
5    synthesis or using a combinatorial synthesis approach.

Alternatively, a PBU may be of unknown sequence content, i.e. the identity of the PBU or any bound biomolecule can be determined retrospectively, if desired, by determining the sequence of both using methods known in the art. A target protein includes proteins and glycoproteins, polypeptides and peptides and also includes protein-nucleic acid complexes.
10   Proteins are generally from a biological source such as a cell, tissue or biological fluid. A biological fluid includes a sample obtained from any site (e.g. blood, plasma, serum, bile, urine, cerebrospinal fluid, aqueous or vitreos humor, or any bodily secretion), a transudate, an exudate (e.g. fluid obtained from an abcess or any other site of infection or inflammation), or fluid obtained from a joint (e.g. a normal joint or a joint affected by a disease such as
15   rheumatoid arthritis, osteoarthritis, gout or septic arthritis), a tissue sample (e.g. a biopsy, blood cells, smears) or homogenates and extracts, including cytoplasm, membranes, and organelles thereof. Cell cultures and culture fluid are also biological samples.

A biological sample may be pre-treated as required. Such a treatment may comprise fractionation, differential extraction (to produce e.g. membrane and cytosolic fractions);
20   selective depletion (e.g. for removal of albumin, haptoglobin, immunoglobin G); and application to any specific affinity column (e.g. mannose-6-phosphate receptor for lysosomal enzymes; Sleat and Lobel, 1997, J. Biol. Chem. 272:731-8).

The method of the invention overcomes the problems of high levels of background label signals inherent in using proteins and especially labelled proteins, and also overcomes
25   issues associated with non-specific binding by using a competitive assay format where specific protein binding to the immobilised PBUs is distinguished from non-specific binding, or from non-specific label adsorption. Thus, binding of a labelled reference sample is measured in the presence of either an unlabelled first sample or a second (or subsequent) unlabelled sample, for example an unlabelled reference or an unlabelled test sample where
30   the reference sample is the first sample and the unlabelled test sample, the second (or subsequent) sample. It is understood that the unlabelled reference sample may alternatively be the second sample and the unlabelled test sample, the first sample. The unlabelled reference and the unlabelled test samples (i.e. an unlabelled first and a second (or subsequent) unlabelled sample) can therefore be compared in their ability to compete for
35   binding to the immobilised PBU with the labelled reference sample. For ease of explanation, which is in no way meant to be limiting, the following text refers to a first sample as an unlabelled reference sample and a second or subsequent sample as a test sample.

When the first sample is a reference sample and the concentration of target protein contained in it is known absolutely, the concentration of the target protein in the test sample
40   may be determined absolutely. If the concentration of target protein contained in the reference sample is not known absolutely then the relative concentration of target protein in the test sample may be determined.

The method described herein substantially compensates for, and can totally eliminate background issues and advantageously no calibration is required.

A labelled reference sample includes any sample used for displacement by an unlabelled first sample or second (or subsequent) samples using the method of the
5      invention. Most preferably, the labelled reference sample is identical to the unlabelled reference sample (i.e. from the same source with components present in the same amounts), but has been labelled for detection purposes. Alternatively, the labelled reference sample and the unlabelled reference samples are non-identical and can be from different biological sources, e.g. different cells or tissues. Preferably samples from different biological
10     sources are from different sources eg different cells or tissues within the same organism. In the most useful application of the method of the invention, the biological sources are sources from humans. The labelled reference sample can also be a pool of protein samples derived from different biological sources. However, it is understood that the disclosed method can employ any labelled reference protein mixture providing said labelled reference sample
15     contains components that bind to all PBUs. This allows the use of generic labelled biomolecular pools, for example but without limitation, obtained by averaging a variety of different cell proteomes (by physical mixing of the protein extracts). This may result in a reduction in the protein concentration range resolved, but significantly, the use of such generic labelled pools would lead to further significant reductions in the time and cost of
20     using arrays of PBUs.

A test sample includes any unlabelled sample that is to be compared with an unlabelled reference sample. It will be understood by one skilled in the art that test samples may also be compared with each other. Most preferably, test samples are from the same biological source such as the same tissue, cell or biological fluid type as the unlabelled
25     reference samples. Alternatively, they are from different biological sources. In one embodiment, the labelled reference sample is derived from more than one biological source. In a particular embodiment, the labelled reference sample is derived from more than one biological source and the unlabelled reference and test samples are from the same biological source.

30     Preferably, samples are clinical samples from, for example but without limitation, diseased versus normal tissue or cell samples where the normal tissue or cell sample is the reference sample and the diseased tissue or cell sample is the test sample. In another preferred embodiment, samples from healthy individuals are compared with samples from diseased individuals and/or with samples from individuals undergoing treatment. It is
35     understood that using the method of the invention permits the comparison of at least two samples on at least two replicate protein binding arrays with no requirement for labelling of said samples for comparison by competing the binding of a labelled reference sample. This is a substantial advantage where the test sample is in short supply, e.g. a biopsy sample. Using unlabelled samples is also particularly advantageous in that this avoids any
40     modification of the proteins present within the sample which can modify or mask interaction sites and which, in turn, may result in the missing of the detection of a potential target of interest.

In the disclosed format, any potential modification of a protein by fluorescent dyes or any other form of labelling will have no deleterious affect on the results as only the displacement of labelled reference proteins by unlabelled proteins present in the reference and test samples are being compared. Thus, even if labelling has modified, for example the

5    antigenic determinants of the proteins in the reference sample, the displacement data is not affected, i.e. the comparison is between the displacement of the labelled reference by the unlabelled reference sample with the displacement of the labelled reference by the unlabelled test sample.

Most preferably, the differential protein composition of two or more samples of

10    interest is assessed such that the displacement of a labelled reference sample e.g. a fluorescently-labelled reference sample, by an identical or similar reference sample that is unlabelled, is compared with the displacement of an identical amount of the same labelled reference sample by an unlabelled test sample. Using an equal amount of unlabelled reference sample to displace the labelled reference sample results in approximately 50%

15    displacement of the labelled reference sample, providing that the concentration used are higher than the relative $K_D$. Similarly, using an equal amount of unlabelled test sample to displace the same said labelled reference sample results in an equivalent displacement of the labelled reference sample unless there is more or less displacement because of a difference in the expression levels of one or more proteins recognised by the immobilised

20    PBUs. Specific protein binding to immobilised PBUs is distinguished from the non-specific binding (which is automatically compensated for, see formulas below) or label adsorption (which can also be due to labelling of for example, immobilised antibody by the fluorescent dye remaining in the labelled reference sample) using this competitive assay format. Thus, generally, if for example, the binding of the 1 μg of a fluorescently-labelled protein present in

25    the labelled sample is competed with an equal amount (1 μg) of the same protein that is not labelled, the decrease in the recorded fluorescence signal will correspond to approximately 50% of the specific binding values, with the rest of the signal being due to non-specific binding. In practice, the displacement values vary slightly due to chemical differences introduced during protein labelling (see Example 2 in the Experimental Section) or if affinity

30    of said binding is low. Displacement using a 100 to 1000-fold excess of unlabelled sample will displace most or all of any specifically bound protein.

Thus, the method of the invention permits the fully quantitative assessment of differential protein composition between samples. The method is particularly useful for assessing differential expression or post translational modification of proteins, and in

35    diagnosis and screening where differential expression or post translational modification of proteins is an indicator of the diseased state.

A quantitative character of the relations between the detected signals and the concentrations of the proteins in the samples studied can be seen from the following formulas. The value of the binding measured (the "Signal") using the described competitive

40    assay can be expressed as:

$$\text{Signal}_{K_i} = B_i + \frac{\text{MAX}}{1 + [K_i]/IC_{50_i}}$$

$$\text{Signal}_{O_i} = B_i + \frac{\text{MAX}}{1 + [O_i]/IC_{50_i}}$$

where: $B_i$ = background signal measurement for each individual PBU.

MAX = maximum signal measurement in the absence of displacement (i.e. signal range).

$IC_{50_i}$ = concentration of an individual protein in each sample, required for a 50% reduction in a signal (is a constant value dependent upon the respective $K_{D_i}$ and [$L_i$].

[$K_i$] = concentration of an individual protein in the unlabelled test sample.

[$O_i$] = concentration of an individual protein in the unlabelled reference sample.

Signal = signal measured for each target protein in the sample ($S_{K_i}$ in the unlabelled test sample; $S_{O_i}$ in the unlabelled reference sample).

[$L_i$] = concentration of an individual protein in the labelled reference sample.

The difference in signal associated with an individual protein between the two unlabelled samples is thus assessed using the following formula:

$$S_{K_i} - S_{O_i} = \frac{\text{MAX}}{1 + [K_i]/IC_{50_i}} - \frac{\text{MAX}}{1 + [O_i]/IC_{50_i}}$$

$$S_{K_i} - S_{O_i} = \text{MAX} \times \frac{(1 + [O_i]/IC_{50_i}) - (1 + [K_i] IC_{50_i})}{(1 + [K_i]/IC_{50_i}) \times (1 + [O_i]/IC_{50_i})}$$

$$S_{K_i} - S_{O_i} = \text{MAX} \times \frac{[O_i]/IC_{50_i} - [K_i]/IC_{50_i}}{(1 + [K_i]/IC_{50_i}) \times (1 + [O_i]/IC_{50_i})}$$

$$S_{K_i} - S_{O_i} = \text{MAX} \times \frac{[O_i] - [K_i]}{(IC_{50} + [K_i]) \times (IC_{50_i} + [O_i])}$$

If [$O_i$] >> $K_{D_i}$ and [$K_i$] >> $K_{D_i}$ where $K_D$ is the dissociation constant for an individual component in the sample, then $IC_{50_i}$ will be approximately equal to [$L_i$] and thus,

$$S_{K_i} - S_{O_i} = \text{MAX} \times \frac{[O_i] - [K_i]}{([L_i] + [K_i]) \times ([L_i] + [O])}$$

further, if both [$O_i$] and [$K_i$] are higher than [$L_i$], i.e. the total concentration of each of [O] and [K] are higher than the total concentration of [L];

for example, and without limitation:

$$[O] \cong 10 \times [L] \text{ and } [K] \cong 10 \times [L];$$

then as in the Example 3 below, the equation can be simplified as follows:

$$\frac{S_{K_i} - S_{O_i}}{MAX} = \frac{[O_i] - [K_i]}{[K_i] \times [O_i]} = \frac{1}{[K_i]} - \frac{1}{[O_i]}$$

$$\frac{1}{[K_i]} - \frac{1}{[O_i]} = \frac{Signal_{K_i} - Signal_{O_i}}{MAX}$$

As noted above, if $[O_i]$ is known absolutely, then $[K_i]$ can be determined absolutely. Otherwise the relationship between $[K_i]$ and $[O_i]$ can be determined from the equation.

The above equation is simplified such that it is not dependent on a background signal measurement (since this is removed on subtraction of $Signal_{O_i}$ from $Signal_{K_i}$), an $IC_{50}$ value or $K_D$, or a maximum signal measurement or even the $[L_i]$. This holds true for each individual protein (i) and permits the method of the invention to be used in a quantitative manner. Since the linear portion of a competitive binding curve spans over two orders of magnitude of a sample concentration a preferred choice of unlabelled reference and test sample concentration (total protein content) is at least 5-fold that of the labelled reference sample concentration, and more preferably at least 10-fold that of the labelled reference sample concentration.

If the above experimental conditions are not met, using the method of the invention gives a qualitative assessment of the differential protein composition between two or more samples, where the higher each individual protein concentration compared to its respective affinity for the PBU, the larger the measured signal difference will be and the signal measured will be inversely proportional to the concentration of respective proteins in the samples.

The lack of dependence of the described method on the reagents used in the production of an affinity array (preparation of which is described below) and on the protein composition of the unlabelled reference and test samples means that the same array format can be used with a variety of sample types. The competitive nature of the disclosed method makes it independent of the protein composition of the unlabelled reference and test samples, unlike the direct binding format of typical array applications where a labelled protein or nucleic acid is bound to an immobilised antibody or nucleic acid or vice-versa and in which the binding signal depends not only on sample concentration, but also on the affinity of the PBU and crossreactivity of the affinity reagents used. In the method of the invention, the displacement achieved by using 1:1 ratio of labelled:unlabelled reference sample will result in approximately 50% displacement regardless of the affinity of the capture reagents or individual protein concentrations in the sample, as long as sample concentrations are kept high ($\gg$ respective $K_D$). Moreover, unlike direct binding formats, where the concentration of the sample must be within a relatively narrow range in order to produce quantitative results, the disclosed method can quantitatively characterise protein concentration in an extremely wide dynamic range. The lower range is limited by the sensitivity of the detection technique (similar to a direct binding approach) but the upper

range is unlimited. Suitable detection techniques include without limitation, fluorescent scanners (for use with fluorescently labelled probes), phospho-imagers (for detecting radioactively labelled samples), MALDI mass spectrometers for use with isotopically or chemically labelled proteins, surface plasmon resonance-based devices for detecting proteins of different sizes. In one embodiment, known signal amplification techniques (sandwich affinity assays, ELISA, rolling circle amplification etc) are used to amplify the differential signals obtained by the described competitive assay. Depending on the detection or the amplification technique used, an appropriate "label" should be selected for labelling the reference sample (i.e. a fluorescent label for direct scanning or an affinity tag or chemically conjugated moiety for labelling "reference" samples for use in sandwhich affinity assays or ELISAs). Such a combination of the competitive assay followed by a signal amplification technique will have all the advantages of the described competitive assay in addition to the increased sensitivity due to signal amplification.

In one embodiment, more than two affinity arrays (each containing a plurality of PBUs) can be used to assess displacement by more than two samples.

In another embodiment, a range of unlabelled reference and test sample amounts is used for displacement of the labelled reference sample such that a displacement curve can be prepared for the assessment of the concentration of the displacing unlabelled reference and test samples of interest.

**Preparation of affinity arrays**

Chosen capture agents are preferably immobilized on a single substrate to form an array of PBUs. Preferably the array is laid out in a grid form at, for example a grid containing 4 or more, preferably 16 or more PBUs. Selection of the solid support and the methods for immobilization depend on the type of capture agents. Immobilization may be performed through covalent or non-covalent binding. In a preferred embodiment, the PBUs are covalently bound to the surface of the substrate via reactive groups or cross-linking agents.

When PBUs are proteins, such as receptors or antibodies, covalent immobilization may be achieved through free cysteines using thiol reactive substrates, or through lysines using amine reactive substrates.

Alternatively, heterobifunctional crosslinkers may be used to cross link -NH groups of the substrate to sulfhydryl (-SH) groups of cysteine residues, e.g. BMPS (beta-maleimidopropionic acid N-hydroxysuccinimide ester) can be used to cross link -NH groups to -SH groups on cysteine residues of antibodies.

Non-covalent interactions may be achieved by using avidin- or streptavidin-coated substrates and biotinylated PBUs; protein-A or protein-G coated substrates and Fc containing PBUs (e.g. immunoglobulins); and metal-chelate substrates and histidine tagged PBUs.

In one preferred method, PBUs are multiple single chain antibodies (SCAs) specific to different peptides. They are purified from phages and solubilised in PBS at similar concentrations (about 0.1mg/ml). The specific SCAs are placed in multiwell plates (e.g. 96-well) and used to generate a microarray system. It is preferable when dispensing proteins to employ a system that does not use steel pins or contact dispensing equipment. Thus, a preferred system is the Packard BCA Piezo robot (US 5,927,547), which dispenses small

volume drops (less than 1nl) from a glass capillary from above the surface of the microarray substrate. A preferred substrate is an aldehyde-activated polyacrylamide pad immobilized on glass or oxidised silicon. A preferred size for the polyacrylamide pad is 2cm x 2cm x 20μm (hydrated depth). About 300pl of each SCA is dispensed into a discrete area within

5   the polyacrylamide pad to create a 2-dimensional array in a 3-dimensional structure. The resulting PBU is preferably about 200μm in diameter. The free aldehydes within the polyacrylamide react with the amines in the SCAs so that the SCAs are covalently immobilized. Once this reaction is complete (e.g. 30 min after dispensing) any remaining aldehydes can be reduced with sodium borohydride or blocked using reagents containing

10  amino groups (e.g. TRIS buffer) and/or free amino acids.

Another preferred substrate is Hydrogel on which PBUs such as antibodies may be immobilized, e.g. via the -NH group of lysine residues to form a Schiff base with aldehyde groups present in the Hydrogel. Alternatively, by soaking Hydrogels in a lysine solution this provides free -NH groups that can be cross-linked to the -SH groups of cysteine residues in

15  antibodies (e.g. via BMPS).

In another preferred method, the substrate is produced by casting a thin agarose gel on a glass slide, using suitable gaskets or other spacers, and siliconised cover glass or cover slip or any similar solid substrate to produce a gel of uniform thickness. Following casting, the gel is activated with CNBr. The use of agarose gel is preferred over commercial

20  Hydrogel (e.g. from Packard Instrument Co., Meriden, CT) as agarose retains moisture more efficiently and thus allows for spotting of dehydration-sensitive reagents. Another advantage of using thin agarose gel films is that they allow for faster diffusion of compounds due to a larger pore size. Therefore, shorter washing steps may be used. This is in contrast to acrylamide based Hydrogel, which require an increased washing time. A further advantage

25  of 3-dimensional gel-based immobilization techniques is that, through careful choice of physical and physicochemical gel properties, it is possible to allow target peptide fragment-PBU binding to closely approximate the binding likely to occur in free solution i.e. less impeded by solid substrates.

In yet another embodiment, single chain Fv antibodies are histidine-tagged. The

30  histidine tag can then be used for immobilization, e.g. using metal chelate affinity on nickel-modified substrates.

Additional immobilization chemistries are possible using other affinity agents and array substrates (e.g. Lin, Science 1997, 278: 840-843).

Other examples of immobilization include using thin membranes (such as

35  nitrocellulose, nylon (charged or uncharged), PVDF and/or their derivatives) attached to or immobilized on glass slides or other solid substrates. For example, attaching commercially available membranes (nitrocellulose, supported nitrocellulose, nylon, charged nylon or PVDF membranes from Schleicher & Schuel, UK Ltd., Amersham, Millipore and other suppliers) to glass slides or by using CAST™ (layer of Nytran® SuPerCharge positively charged nylon

40  membrane affixed to the glass) and FAST™ slides (microporous polymeric substrate cast onto glass from Schleicher & Schuel, UK, Ltd) may lower costs significantly. For this method, PBUs may be dispensed using piezo robots similar to the one described above for SCAs. One of the advantages of using nitrocellulose or nylon based supports is that they

have a much higher protein binding capacity compared to Hydrogel or 2 dimensional solid supports (glass, silicon, plastic, etc.). Various solid support materials that may be used in the microarray are as follows:

Binding capacity (highest to lowest):

5 a) immobilized nitrocellulose; immobilized nylon (charged > uncharged)
  b) FAST; CAST
  c) Hydrogel
  d) derivatized solid substrates (glass, silicon, plastics)

10 PBU geometry (best to worst):
  a) derivatized solid substrates (glass, silicon, plastics)
  b) FAST slides/CAST slides
  c) immobilized nitrocellulose; immobilized nylons

15 Variability in spotting (least variable to most variable):
  a) FAST slides
  b) CAST slides
  c) immobilized nitrocellulose
  d) immobilized nylons
20 e) Hydrogels

Moisture retaining (best to worst):
  a) agarose pads
  b) acrylamide pads (Hydrogel)
25 c) immobilized membranes or derivatized solid substrates (glass, silicon, plastics)

Tolerance to glycerol in the spotting mixtures (highest to lowest):
  a) FAST slides; immobilized nitrocellulose
  b) CAST slides; immobilized nylons
30

### Hydrogel

Another preferred immobilization technique uses a combination of non-covalent affinity immobilization coupled to covalent cross-linking. This technique is especially advantageous for immobilizing PBUs such as antibodies, as it allows the antibodies to be

35 attached in an orientation whereby their active sites are easily accessible, and at the same time, remain covalently cross-linked to the solid support. This technique includes, but is not limited to, using Hydrogel (Packard Instrument Co., Meriden, CT) for immobilization of protein-A or protein-G, followed by treatment with glutaraldehyde (high concentration for short time). This results in covalent cross-linking of the protein-A/G to the solid support, and

40 also in aldehyde derivation of solution-exposed amino groups of the protein-A/G. Substrates are then washed to remove the excess free glutaraldehyde, while bound aldehyde groups remain un-blocked. Subsequently, Fc containing antibodies (e.g. IgG) are spotted or dispensed onto the derivatized protein-A/G, which then binds the Fc fragments of the antibodies. Following incubation, protein-A/G eventually cross-links with the bound

45 antibodies through aldehyde-derivatized groups on the protein-A/G and available amino

groups on the Fc fragments. This approach results in high density antibody immobilization, with antibodies being correctly oriented and covalently cross-linked to the solid support. Glutaraldehyde pre-treatment of the immobilized Protein-A/G prior to antibody spotting is a preferred technique because it avoids incubating the PBUs (i.e., the antibodies) with the
5    cross-linking agent (i.e. glutaraldehyde), which could result in the derivatization of amino groups of the antibodies.

It is well known that many proteins contain nucleic acid-binding domains. As such, the use of nucleic acids as PBUs, e.g. as nucleic acid arrays, will allow the capture of proteins containing nucleic acid-binding domains. Nucleic acid arrays may be formatted as
10    "gene chips" and related arrays of oligonucleotides, cDNAs, and other nucleic acids, which are well known in the art (see for example the following: US 6,045,996; 6,040,138; 6,027,880; 6,020,135; 5,968,740; 5,959,098; 5,945,334; 5,885,837; 5,874,219; 5,861,242; 5,843,655; 5,837,832; 5,677,195 and 5,593,839.

15    **Sample Labelling**

Samples may be labelled according to any standard method known in the art to facilitate detection. Preferred methods include the use of Fluorophore-iodoacetamide, -maleimide or -succinimide groups (e.g. Molecular Probes) Cy3, Cy5 or TAMRA dyes for example via labeling cysteines or lysines (e.g. Alexa Fluor dyes). Another preferred method
20    makes use of NHS ester Bodipy TMR (Molecular Probes, Leiden, Netherlands). Samples may also be labelled via aldehyde or ketone groups (which may be selectively introduced), using hydrazine dyes (Molecular Probes). Other labels for use in the invention include but are not limited to fluorescein isothiocyanate (FITC), phycoerythrin (PE), Texas red (TR), rhodamine, free or chelated lanthanide series salts, especially $Eu^{3+}$, chemiluminescent
25    molecules, radio-isotopes ($^{125}I$, $^{32}P$, $^{35}S$, chelated Tc) or magnetic resonance imaging labels.

Metabolic (or biosynthetic) labeling with [$^{35}S$]-methionine, is also contemplated as well as labeling with [$^{14}C$]-amino acids and [$^{3}H$]-amino acids (with the tritium substituted at non-labile positions).

Other aspects of the invention claimed herein include use of the method as
30    hereinbefore described in determining differential expression or post translational modification in proteins between two or more samples containing target proteins.

There is also provided the method for use in the screening, diagnosis or prognosis of a disease in a subject, for determining the stage or severity of said disease in a subject, for identifying a subject at risk of developing said disease, or for monitoring the effect of therapy
35    administered to a subject suffering from said disease

According to the method for use in the screening, diagnosis or prognosis of a disease in a test subject, for determining the stage or severity of said disease in a test subject, for identifying a subject at risk of developing said disease, or for monitoring the effect of therapy administered to a test subject suffering from said disease, one of a first, second and
40    subsequent samples is a sample from a non-diseased subject, one is a sample from a diseased subject and one is a sample from the test subject.

In the aforementioned method, generally the method will be performed in order to assess if levels of characteristic proteins are raised or lowered, such raising or lowering known to be characteristic of the diseased state, or characteristic of prognosis. The levels can also be monitored in a subject under therapy to assess the response to therapy.   Those proteins the raising or lowering of levels of which are characteristic of the diseased state are either or known or may be determined by conventional methods.

There is also provided a  kit for use in a method (especially a diagnostic kit for diagnosing disease) as hereinbefore described which comprises:

    (a)    a labelled reference sample;

    (b)    an unlabelled first sample from a non-diseased subject serving as a negative control;

    (c)    an unlabelled second sample from a diseased subject serving as a positive control; and

    (d)    replicate substrates having immobilised upon them one or more protein binding units, differential binding to which is capable of characterising the presence or absence of disease in the test subject.

The replicate substrates are used in the steps of the method: one in step (a) for use with the unlabelled first sample and two or more in step (b) for use with the second sample and the sample(s) from the test subject(s). Differential binding of the samples to the replicate substrates is capable of indicating relative levels of characteristic proteins as discussed earlier in the description.

There is provided in particular a diagnostic kit for use in a method of diagnosing disease, which disease is characterised by a raised or lowered level of characteristic target proteins in a sample from the test subject.

The method of the invention has significant advantages over current technologies, viz:

1. labelling of the samples for comparison is not required.

2. the invention permits the comparison of unlabelled samples, thus avoiding inaccuracies due to sample modifications produced by the introduction of a label.

3. the use of non-labelled samples means that no purification step is required resulting in faster and cheaper experiments.

4. generic arrays, protocols and detection means can be used for any sample type, regardless of its protein composition, thus enabling mass production that is standard and cost effective.

5. the method is especially suitable for protein samples that are in limited supply such as, but without limitation, biopsy samples.  A single labelled reference sample, e.g. an appropriate control or untreated sample, can be prepared on a large scale for use in all experiments, with different unlabelled test samples used for displacement. The method also allows the use of a generic protein mixture for labelling thus further reducing the cost and time spend for each experiment.

6. a broad dynamic range of protein concentrations is resolvable with no upper limit.

7. the effect of background non-specific adsorption is removed by using displacement data making it possible to utilise a variety of substrates.

8. the method does not depend on the affinity of the capture reagents used.

5

The method is equally applicable to a variety of assays such as but not limited to, protein-protein interactions, antibody-protein interactions, and protein-nucleic acid interaction pairs.

All publications, including, but not limited to, patents and patent applications, cited in this

10    specification, are herein incorporated by reference as if each individual publication were specifically and individually indicated to be incorporated by reference herein as though fully set forth.

The invention will now be described by reference to the following examples and figures which are merely illustrative and are not to be construed as a limitation of the scope of the

15    present invention.

Description of the Figures

Figure 1 is a map showing the positions of substrate-immobilised PBUs comprising capture agents in an affinity array of 74 PBUs.

20

Figure 2 is a scanned image of five arrays labelled I-V, showing the degree of fluorescent labelling of both immobilised antibodies and of the background acrylamide gel obtained using five fluorescent dyes (see Example 1a).

25    Figure 3 is a series of histograms showing the degree of fluorescent labelling of the immobilised antibodies (black fill) and of the background acrylamide gel (no fill) obtained using five fluorescent dyes (see Example 1a). The histogram stacks labelled I-V correspond to those in Figure 2.

30    Figure 4 is a series of histograms showing the displacement of fluorescently-labelled reference whole liver lysate by various amounts (indicated in panel D) of an identical unlabelled reference whole liver lysate on an array comprising PBUs (see Example 1b). Panel A relates to row A, columns 1-19 of Figure 1; Panel B relates to row B, columns 1-19 of figure 1; Panel C relates to row C, columns 1-19 of Figure 1; Panel D relates to row D,

35    columns 1-19 of Figure 1. Each histogram is composed of five stacks, each one representing one displacement amount indicated on the left hand side of Panel D; Stack 1 = 0µg and is always close to maximal; stack 2 = 0.1µg; stack 3 = 1µg; stack 4 = 10µg and stack 5 = 100µg and is always minimal being approximately zero%.

40    Figure 5 shows a series of histograms of the results obtained from an array of substrate-immobilised PBUs comprising 4x18 antibodies contacted with labelled ovary lysate in the presence of unlabelled ovary lysate (unlabelled reference sample) or unlabelled kidney lysate (test sample) as described in Example 2. Filled bars represent the difference between the unlabelled reference sample and unlabelled test sample (ovary sample minus kidney

sample). Each panel (A-D) corresponds to a single row of PBUs (see Table 1). Open bars in each case show mean differences obtained by averaging data from four PBUs (each PBU was printed in quadruplicate). Filled bars show mean differences obtained by averaging data from two PBUs (i.e. two outlying values were disregarded in each case).

Figure 6 is a repetition of the experiment described in Figure 5 except that the test sample (kidney lysate) was used in x10 excess, i.e a 1:10 ratio (by total protein) compared to the labelled reference sample, whilst the unlabelled reference sample (ovary lysate) was used at 1:1 ratio with the labelled reference sample (see Example 3). Panels A-D and open and filled bars are as in Figure 5.

## Example 1(a).
### Selection of Label

To assess background levels of fluorescence obtained with different dyes due to non-specific dye binding-adsorption, silicon slides with polyacrylamide gel pads (generated as described above or obtained from Packard Bioscience, Meriden, CT, USA) were treated with glutaraldehyde (10-50%) overnight (optionally followed by 2% TFA for 5 min), extensively washed with water, and air-dried. PBUs were prepared by dispensing capture agents using an arrayer (BioChip Arrayer with Piezo-tips, Packard Bioscience, Meriden, CT). To prepare PBUs, slides were incubated with the dispensed capture agent in a sealed, humidified chamber overnight. To block the remaining reactive sites, and to remove unbound capture agents, slides were washed in Tris-Buffered Saline (0.05M Tris, 0.15M NaCl, pH7.6):glycine (1.5% w/v in water):bovine serum albumin (1% w/v final), followed by a final rinse in water before drying. The slides comprising the PBUs were then ready for use.

### Incubation of fluorescent dyes with antibody arrays

The fluorescent dyes detailed below were prepared (at equivalent concentrations) in 1 x Tris-Buffered Saline plus bovine serum albumin (1% w/v final). Each dye was incubated with one of five identical arrays (I-V) prepared as described above at room temperature for 1hr. Arrays were washed in $H_2O$ for 15min with several changes of water. The identities and positions of each PBU are shown in Figure 1 (antibodies -columns 1-10, from Sigma, Chemical Co. USA, and transcription factors -columns 11-19; Abcam, Ltd, Cambridge, UK).

Fluorescent dyes:
1. Iodoacetamide-Cyanine5 (-Cy5) – Array I
2. Iodoacetamide-Cyanine3 (-Cy3) – Array II
3. Succinimide-Alexa-633 – Array III
4. Maleimide-Alexa-633 – Array IV
5. Succinimide-Tamra – Array V

The images of the five arrays incubated with fluorescent dyes are shown in Figure 2 corresponding to Panels I-V. The incorporation of fluorescent dye molecules into immobilised antibodies and acrylamide gel was measured using a Biochip Imager Confocal laser scanning system (Packard BioScience, Meriden, CT, USA) at two wavelengths: 543

nm (for iodoacetamide-Cy3 and succinimide-Tamra dyes) and 633 nm (for iodoacetamide-Cy5, succinimide-Alexa-633 and maleimide-Alexa-633 dyes). The images were all scanned at the highest resolution of 10µm. The results are shown in Figure 3.

5      For the arrayed antibodies the highest signal intensities were observed with iodoacetamide-Cy5, maleimide-Alexa-633 and succinimide-Alexa-633 dyes whereas much lower incorporation was seen using iodoacetamide-Cy3 and succinimide-Tamra dyes. The incorporation of dye into the gel was at a low level for all dyes examined except for iodoacetamide-Cy5 where a high background level was exhibited. A good signal (i.e. dye incorporation into protein) to background (i.e. very little dye remaining in acrylamide gel

10     following washing) ratio was achieved with both the Alexa dyes suggesting they are the superior dyes for use with this system although all dyes tested are suitable.

**Example 1(b).**

       Whole liver lysate was labelled with iodoacetamide-Cy5 and used as the labelled
15     reference sample.

*Labelling Protein Lysates*

       Protein lysates (1mg) were labelled using an excess concentration of iodoacetamide cyanine-5 dye in phosphate buffered saline buffer, pH7.4, plus Tween (0.5% v/v) at room temperature for 3hr. Following labelling, protein lysates were purified from unconjugated dye

20     by gel-filtration using p-6 chromatography columns (BioRad). Samples were centrifuged at 1,000 x g for 5min to isolate labelled proteins from free, unconjugated dye (samples passed through 2 separate columns). The incorporation of iodoacetamide Cy5 into protein lysates and the removal of 'free' dye was evaluated by examining the purified labelled-protein samples on denaturing acrylamide gels and fluorescent scanning (Biochip Imager Confocal

25     laser scanning system).

       Five identical arrays represented in Figure 4 as individual stacks within each histogram (74 PBUs on each array as detailed in Figure 1) were incubated with labelled reference whole liver lysate (1µg protein) alone or with the labelled reference lysate (1µg of protein) plus an identical sample of unlabelled lysate containing 0µg, 0.1µg or 1µg or 10µg

30     or 100µg protein (see Figure 4). Fluorescent signals were measured in each case for all PBUs as described in Example 1(a) (see Figure 1 for PBU identities). Fluorescence measured after hybridisation in the absence (0µg) of unlabelled protein includes both non-specific binding/adsorption and specific binding components and is maximal, whereas fluorescence measured after hybridisation of 1µg of labelled reference whole liver lysate in

35     the presence of 100µg of unlabelled whole liver lysate contains only non-specific binding/adsorption (and is minimal) and was subtracted from all the experimental data. A maximum of the two fluorescence signal values measured for each individual PBU for the array hybridised with 1µg labelled reference whole liver lysate in either (i) the absence of unlabelled whole liver lysate or (ii) the presence of 0.1µg of unlabelled whole liver lysate,

40     minus the value measured after hybridisation in the presence of 100µg of unlabelled whole liver lysate (non-specific binding/sorption), were taken as maximum (100%) for each individual PBU.

Fluorescent signals for each array minus values measured after hybridisation in the presence of 100µg of unlabelled protein were expressed as % of the maximum (see above) and are shown in Figure 4. The mean displacement for 1µg labelled with 1µg unlabelled sample was 49.8% +/- 2.7% (for an array of 74 PBUs). These results exemplify the

5      quantitative character of the assay (the theoretical expected value is 50% for 1µg displacement) as well as the reliability of the disclosed data extraction technique, which is unaffected by non-specific background.

**Example 2.**

10     _Preparation of substrate-immobilised PBUs_

Hydrogel glass slides (Packard Bioscience, Meriden, Connecticut, USA) were soaked overnight in 50% glutaraldehyde, washed with $H_2O$ and air-dried. To prepare the PBUs, the antibodies (capture agents) were dispensed onto the Hydrogel slides (acrylamide-based pads) using a BioChip Arrayer with Piezo-electric tips (Packard Bioscience, Meriden,

15     Connecticut, USA). A total of 72 antibodies were arrayed in quadruplicate onto each slide (see Table 1). Slides were incubated overnight in a humidified chamber prior to washing by Tris-Buffered Saline (0.05M Tris, 0.15M NaCl, pH7.6):glycine (1.5% w/v in water):bovine serum albumin (1% w/v final) at RT for 2 hours, followed by a final rinse in water before drying. Protein extracts of ovary and kidney were obtained from Clontech. Ovary sample

20     was labelled with iodoacetamide-cy5 and purified using gel filtration columns for use as the labelled reference sample. The equivalent of 1µg of the labelled ovary sample was used for each individual hybridisation. Unlabelled proteins (reference ovary or test kidney) were added as 10µg protein, i.e. each lysate was used in x10 excess (by total protein) compared to the labelled ovary (reference) sample. Slides were contacted with sample for 1hr at RT

25     followed by three 2min washes with $H_2O$.

The binding of fluorescent dye molecules onto the slides was measured using a Biochip Imager Confocal laser scanning system at the highest resolution of 10µm. Total fluorescence signals corresponding to each PBU (obtained by hybridisation in the presence of various displacing samples) were recorded. Mean values (obtained by averaging the

30     quadruplicate readings for each antibody) were used for analysis. The set of values corresponding to the hybridisation in the presence of unlabelled kidney sample was subtracted from an analogous set obtained when unlabelled ovary was used. When the subtraction results in a positive value, this indicates that the signal measured for the unlabelled ovary sample is stronger, i.e. less labelled ovary displaced; thus a higher

35     concentration of the corresponding protein exists in the kidney sample. A negative value means lower protein expression in the kidney sample.

Table 1.

| Panel | | | |
|---|---|---|---|
| A1 | α-2- Macroglobulin | C1 | BCMP11 |
| A2 | Glutathione-S-Transferase | C2 | c-myc |

| A3 | Complement C4 | C3 | P38MAPK |
|---|---|---|---|
| A4 | HLA-class I antigen | C4 | Neurofilment 200 (Ab2) |
| A5 | Actin | C5 | N-CAM |
| A6 | Blood Brain Barrier | C6 | Phosphothreonine |
| A7 | Human Factor IX -5 | C7 | Human Factor IX -1 |
| A8 | Cytokeratin (6B10) | C8 | OGS 549.13 -1 |
| A9 | OGS 549.2 -1 | C9 | OGS 549.4 -1 |
| A10 | OGS 549.26 -1 | C10 | OGS 549.5 -1 |
| A11 | OGS 549.5 -3 | C11 | OGS 549.10 -2 |
| A12 | OGS 549.9 -2 | C12 | E2F2 |
| A13 | GFI-1 | C13 | IkB alpha |
| A14 | MAD | C14 | Nut77 |
| A15 | SP-1 | C15 | TATA binding protein |
| A16 | HNF-1 | C16 | NAT1 |
| A17 | Transcription Enhancer Factor -1 | C17 | ZBP89 |
| A18 | Heparanase (Ab4) | | |
| | | | |
| B1 | BCMP84 | D2 | Bcf81 |
| B2 | Poly-Histidine | D3 | Thrombin |
| B3 | Proliferating Cell Nuclear Antigen | D4 | Matrix Metalloproteinase -3 |
| B4 | Neurofilament 200 (Ab1) | D5 | HSP70 |
| B5 | β-Catenin | D6 | Glial Fibrillary Acidic Protein |
| B6 | Phosphoserine | D7 | Phosphotyrosine |
| B7 | Human Factor IX | D8 | Cytokeratin 5 |
| B8 | Cytokeratin 4.62 | D9 | OGS 549.13 -2 |
| B9 | OGS 549.2 -2 | D10 | OGS 549.4 -2 |
| B10 | OGS 549.26 -1 | D11 | OGS 549.5 -2 |
| B11 | OGS 549.10 -1 | D12 | OGS 549.9 -1 |
| B12 | E2F1 | D13 | Fos |
| B13 | TFIIB | D14 | Jun |
| B14 | NF-YA | D15 | SC35 |
| B15 | Sterol-regulatory element binding protein-1 | D16 | Ets-1 |

| B16 | HSF-4 | D17 | TAF172 |
|-----|-------|-----|--------|
| B17 | S11 (TFIIS) | D18 | Heparanase (Ab3) |
| B18 | Heparanase (Ab8) | | |

Figure 5 shows displacement data obtained for an array of 72 PBUs, incubated and bound with 1µg labelled ovary lysate and displaced by either 10µg unlabelled ovary lysate or by 10µg unlabelled kidney lysate. The plots (Figures 5A-D) show the differences (ovary displacement minus kidney displacement) obtained at approximately 90% overall displacement level (e.g. 10-fold excess of the unlabelled sample, by total protein concentration). Normalised data were obtained as described previously. In all cases, at least one array is required for each test sample. Importantly, regardless of the format (whether a qualitative or partially or fully quantitative analysis), the displacement protocols described herein result in reliable and meaningful data, unlike a direct (non-competitive) hybridisation approach, which works best with nucleic acids.

The results shown in Figure 5 indicate a number of differentially expressed proteins. The highest differential expression was detected using anti-HLA-class I antigen antibody, anti-blood-brain barrier (neurothelin) HT-7 antibody, anti-MAD (helix loop helix - leucine/zipper transcription factor) antibody, anti-phosphothreonine antibody, anti-IkB alpha antibody and anti-TATA binding protein antibody. The results are in a good agreement with data reported previously. For example, reports by Le, et al. (2002, Exp. Mol. Med. 34:18-26) and Dolo, et al. (1997, Oncol Res 9:129-38) indicate expression of HLA-1 in ovary reaching especially high levels in ovarian tumours. The negative value of the HLA-1 (Figure 5A, no.4) is indicative of higher HLA-1 content in the ovary sample. The blood brain barrier protein, neurothelin, recognised by the HT-7 antibody, is predominantly expressed by brain endothelial cells. It has also been reported to be expressed in basolateral membranes of kidney tubules (Risau, W. et al., 1986, EMBO J. 5:3179-8313). As such, a higher level of this protein was found in kidney lysate compared to ovary (Figure 5A, no.6). Anti-phosphothreonine antibody signal (Figure 5C, no.6) is indicative of a higher level of phosphorylated proteins in kidney. MAD protein expression is higher in ovary than kidney (Figure 5A, no.14). The expression profile of IkB alpha protein (Figure 5C, peak 13) shows a higher level of IkB alpha in kidney compared to ovary. The TATA-binding protein (TBP) is a transcription factor, which plays an important role in eukaryotic gene expression and is present in both ovary and kidney on mRNA level (UniGene) shows a higher level of expression in kidney (Figure 5C, peak 15).

**Example 3**

This example demonstrates the tolerance of the method of the invention to variability in the protein concentration of unlabelled reference and test samples. Thus, the same experiment as described in Example 2 was performed except that an array of 72 PBUs was contacted with 1µg of labelled ovary lysate in the presence of either 1µg unlabelled ovary lysate or by 10µg unlabelled kidney lysate (see Figure 6). To compensate for differences in displacement values due to different amounts of protein used for displacement of individual

samples, which can result in a strong bias towards one or another sample, a complete set of subtracted values was averaged and the obtained average value was subtracted from each individual peak. Such normalisation is able to compensate for the overall differences in the amount of displacing reagents if they are within two log units of their respective $IC_{50}s$. Figure

5 6 demonstrates substantially similar results to those shown in Figure 5, with differential expression detected using anti-HLA-class I antigen antibody, anti-blood-brain barrier (neurothelin) HT-7 antibody, anti-MAD (helix loop helix - leucine/zipper transcription factor) antibody, anti-phosphothreonine antibody, anti-IkB alpha antibody and anti-TATA binding protein antibody. Thus, HLA-1 (Figure 6A, no. 4) and MAD protein (figure 6A, no. 14) are

10 higher in ovary than kidney whilst HT-7 (Figure 6A, no.6), threonine phosphorylated proteins (Figure 6C, no.6), IkB (Figure 6C, no.13) and TATA binding protein (TBP; Figure 6C, no.15) demonstrate higher expression in the kidney lysate. The six most differentially expressed signals correspond to the same antibodies as described in the Figure 5 legend, which confirms the high tolerance of the reported method to variability in protein concentrations.

15 The results of Examples 2 and 3 show the same differential expression profile discussed above and demonstrate the quantitative nature of the method of the invention; fold differences can be calculated from the data shown.

**CLAIMS**

1.      A method of comparing the level of a target protein within a first sample with the level of said target protein within a second or subsequent sample, the method comprising:

(a)      contacting a labelled reference sample containing said target protein in the presence of the first sample with an immobilised protein binding unit capable of capturing said target protein;

(b)      for each second or subsequent sample contacting a labelled reference sample which is a replicate of that of step (a) in the presence of the second or subsequent sample with a immobilised protein binding unit which is a replicate of that of step (a); and

(c)      detecting the level of label present on the protein binding units employed in steps (a) and (b) and from a comparison of said levels, determining a comparison of said target protein levels within the first sample and within the second or subsequent sample.

2.      The method according to claim 1, wherein the total concentration of protein in the first and second or subsequent samples of parts (a) and (b) is greater than the total amount of protein in the labelled reference sample.

3       The method according to claims 1 or 2, wherein the total concentration of protein in the first and second or subsequent samples of parts (a) and (b) is at least 5 fold that of the total concentration of protein in the labelled reference sample.

4.      The method according to any one of claims 1-3, wherein the labelled reference and first and second or subsequent sample(s) of parts (a) and (b) are all derived from the same biological source.

5.      The method of any one of claims 1-3, wherein the unlabelled reference and first sample of part (a) are derived from a different biological source within the same organism to the second or subsequent sample of part (b).

6.      The method of any one of claims 1-3, wherein the unlabelled reference sample of part (a) is derived from a different biological source within the same organism to the first sample of part (a) and the second or subsequent sample(s) of part (b).

7.      The method of any one of claims 1-6, wherein the labelled reference sample is derived from more than one biological source.

8.      The method according to any one of claim 1 to 7 wherein in step (c) the level of label present on the protein binding units employed in steps (a) and (b) is detected as a signal in a detection experiment and the difference between the signal obtained in step (a) and that obtained in step (b) for each second or subsequent sample is calculated in order to determine a comparison of said target protein levels within the first sample and within the second or subsequent sample.

9.    The method according to any one of claims 1-8, wherein in steps (a) and (b) the samples are contacted with a plurality of different immobilised protein binding units and in step (c) the level of label present on each different protein binding unit is detected.

5

10.    The method according to claim 9 wherein the plurality of different protein binding units are immobilised on a single substrate.

11.    A method according to any one of claims 1-10, wherein part (b) is performed for a
10    second and at least one subsequent sample.

12.    A method according to any one of claims 1-11, wherein one of the first sample and the second or subsequent samples is a negative control and another is a positive control.

15    13.    The method according to any one of claims 1 to 12 for use in determining differential expression or post translational modification in proteins between two or more samples containing target proteins.

14.    The method according to any one of claims 1-13, for use in the screening, diagnosis
20    or prognosis of a disease in a subject, for determining the stage or severity of said disease in a subject, for identifying a subject at risk of developing said disease, or for monitoring the effect of therapy administered to a subject suffering from said disease.

15.    A method according to claim 14 for use in the screening, diagnosis or prognosis of a
25    disease in a test subject, for determining the stage or severity of said disease in a test subject, for identifying a subject at risk of developing said disease, or for monitoring the effect of therapy administered to a test subject suffering from said disease, wherein one of a first, second and subsequent samples is a sample from a non-diseased subject, one is a sample from a diseased subject and one is a sample from the test subject.
30

16.    A kit for use in a method according to claim 14 or claim 15, which comprises:
      (a)    a labelled reference sample;
      (b)    an unlabelled first sample from a non-diseased subject serving as a negative control;
35    (c)    an unlabelled second sample from a diseased subject serving as a positive control; and
      (d)    replicate substrates having immobilised upon them one or more protein binding units, differential binding to which is capable of characterising the presence or absence of disease in the test subject.
40

17. A kit according to claim 16, wherein in part (d) the replicate substrates have immobilised upon them four or more protein binding units.

18. A kit according to claim 16 or 17, which is a diagnostic kit for use in a method of diagnosing disease, which disease is characterised by a raised or lowered level of characteristic target proteins in a sample from the test subject.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A | A1AT | Macrogl | GST | C4 | HLA1 | Actin | BBB | HIX5 | CK6B10 | 549.2-1 | 549.26-1 | 549.5-3 | 549.9-2 | Gli1 | MAD | SP-1 | HNF-1 | TEF1 | Hepar-4 |
| B | Transte m | BCMP-84 | PolyHis | PCNA | NF200-1 | Catenin | Phosph oserine | HFIX | CK4.62 | 549.2-2 | 549.26-2 | 549.10-1 | E2F1 | TFIIB | NF-YA | SREBP-1 | HNF-4 | S11 | Hepar-8 |
| C | Haptogl | BCMP-11 | NGF | P38MA PK | NF200-2 | N-CAM | Phosph o-Thr | HIX1 | 549.13-1 | 549.4-1 | 549.5-1 | 549.10-2 | E2F2 | IkB-alpha | Nut77 | TATA-BP | NAT1 | ZBP89 | |
| D | Albumin | Bcl81 | Thmbin | BCMP-11-neat | Hsp70 | GFAP | Ph-tyrosine | CK5 | 549.13-2 | 549.4-2 | 549.5-2 | 549.9-1 | fos | jun | SC3S | Ets-1 | TAF172 | Hepar-3 | |

Figure 1

| | |
|---|---|
| I | Iodoacetamide Cy5 |
| II | Iodoacetamide Cy3 |
| III | Maleimide Alexa-633 |
| IV | Succinimide Alexa-633 |
| V | Succinimide Tamra |

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
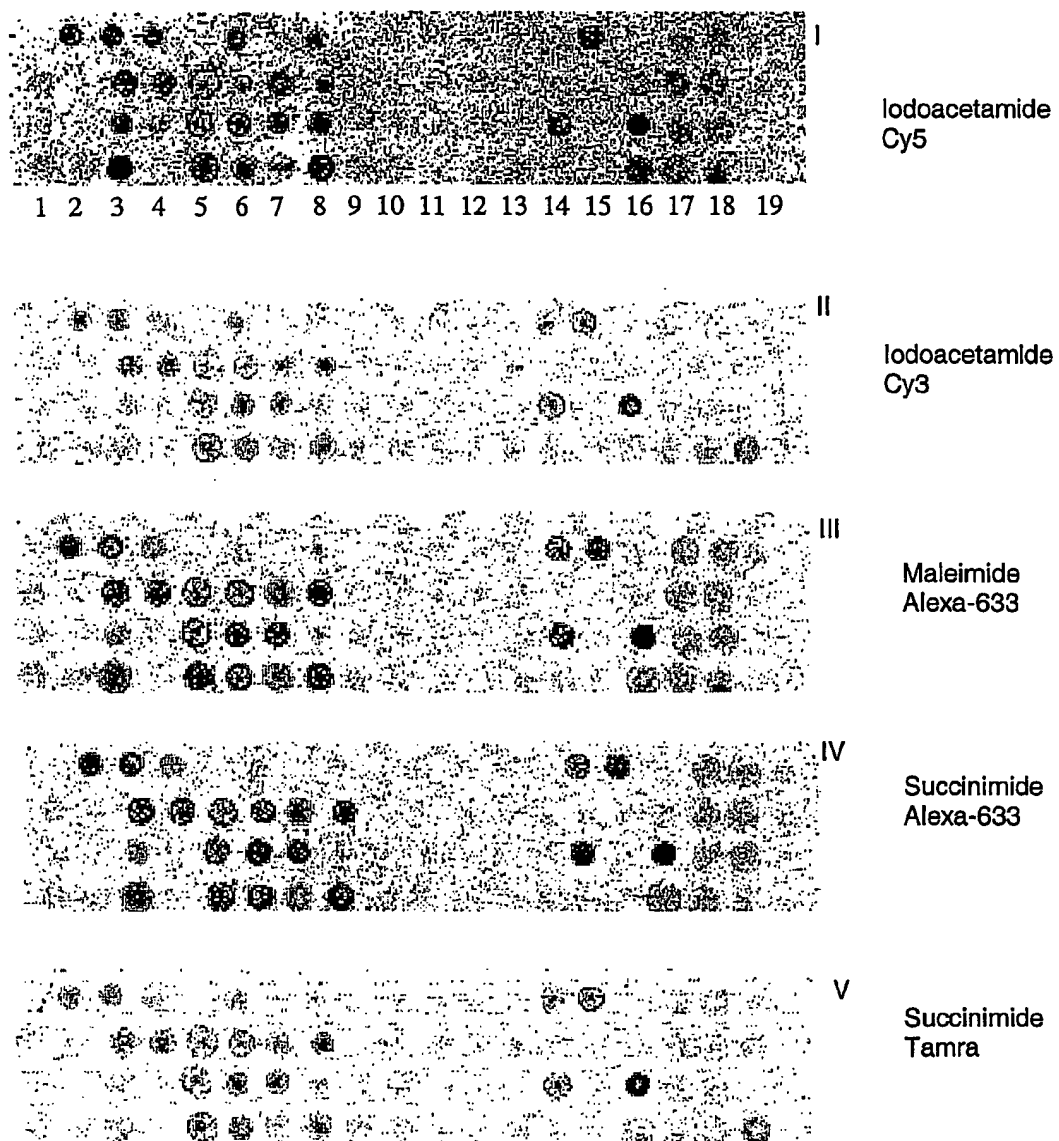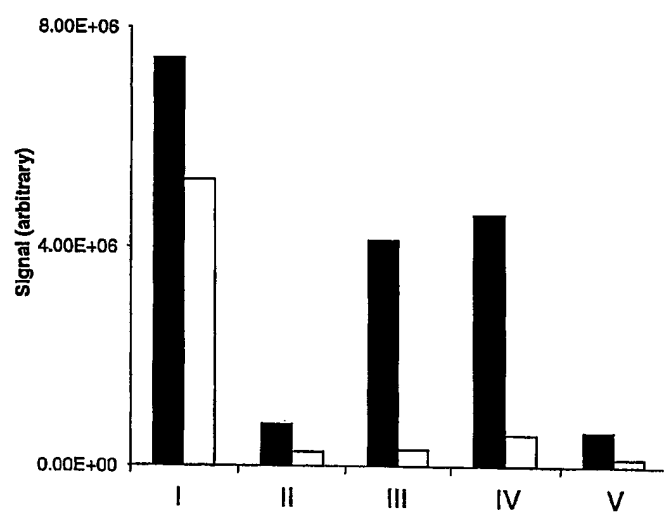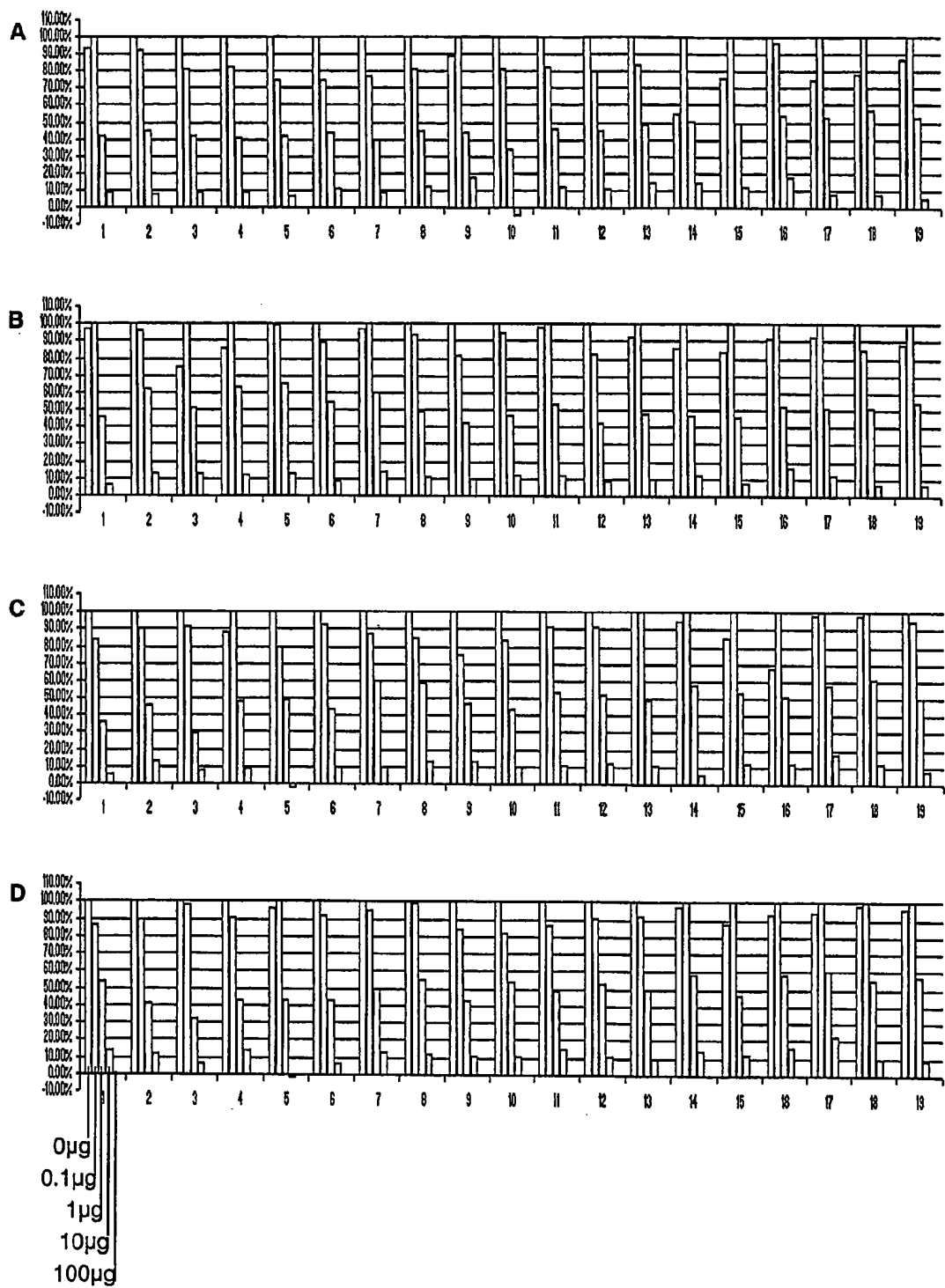
Figure 2

Figure 3

Figure 4
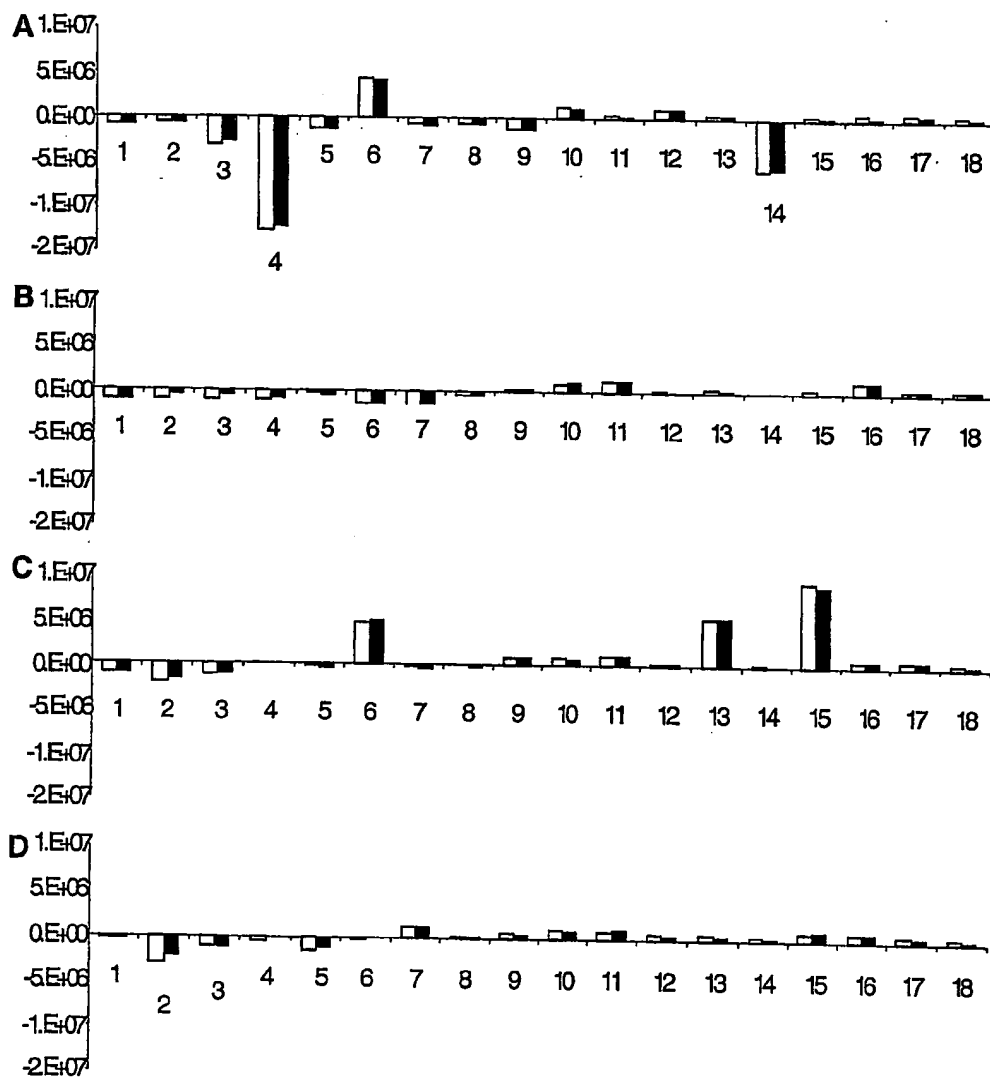
Figure 5
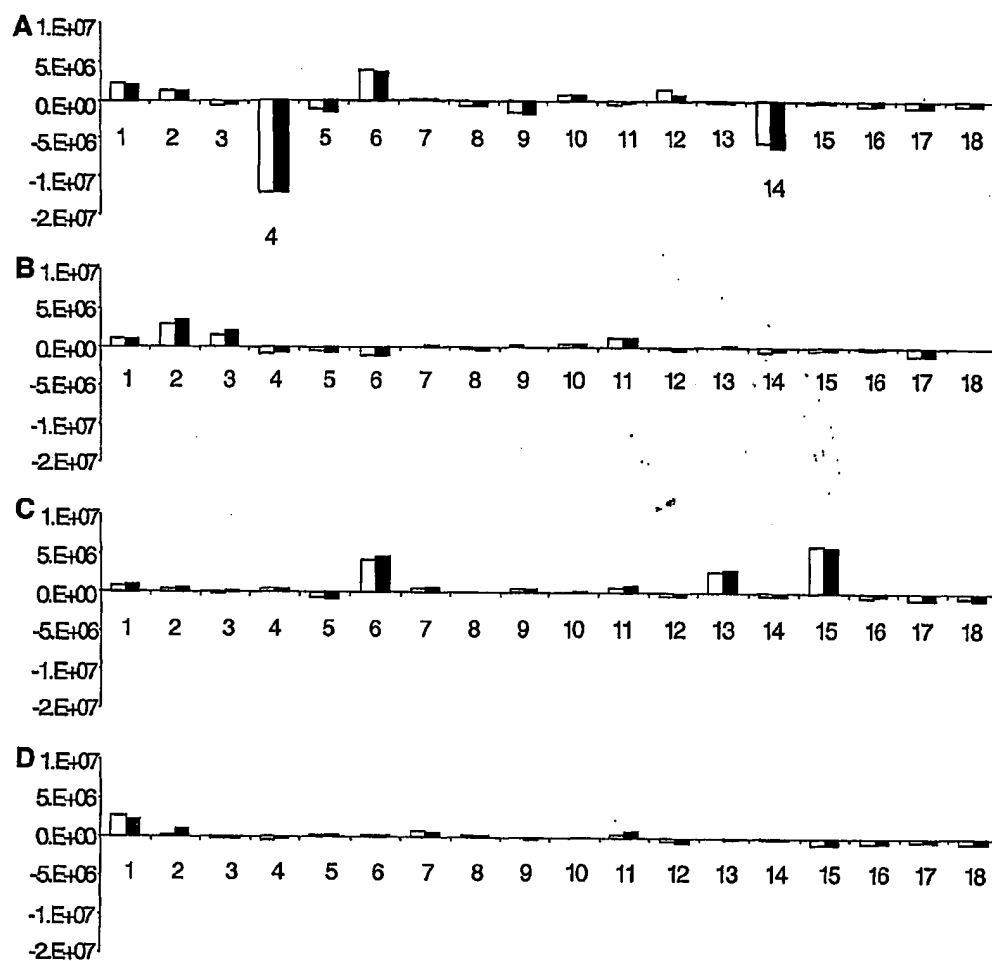
Figure 6

# INTERNATIONAL SEARCH REPORT

## A. CLASSIFICATION OF SUBJECT MATTER
IPC 7    G01N33/68

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched  (classification system followed by classification symbols)
IPC 7    G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category * | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| Y | WO 98 24935 A (AN GANG ;HARA MARK O (US); RALPH DAVID (US); VELTRI ROBERT (US); U) 11 June 1998 (1998-06-11) page 48, paragraphs 3,4; claims 10,24,26,27 page 50, paragraph 4 --- | 1-18 |
| Y | US 4 895 809 A (SCHLABACH TIMOTHY D  ET AL) 23 January 1990 (1990-01-23) cited in the application abstract; claims 1-8 --- | 1-18 |
| A | US 6 197 599 B1 (WANG YINGYI  ET AL) 6 March 2001 (2001-03-06) cited in the application column 3, paragraphs 7,9 column 5, paragraph 2; example 3 --- | 1-18 |
| | -/-- | |

[X]  Further documents are listed in the continuation of box C.      [X]  Patent family members are listed in annex.

**C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT**

| Category * | Citation of document, with indication,where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| P,X | WO 02 10338 A (GETZENBERG ROBERT H ;GENE LOGIC INC (US); KULKARNI PRAKASH (US); M) 7 February 2002 (2002-02-07) page 2, paragraph 4 page 22, paragraphs 2,3 page 28, paragraph 2 -page 29, paragraph 1 ----- | 1-18 |

# INTERNATIONAL SEARCH REPORT

| Patent document cited in search report | | Publication date | Patent family member(s) | | Publication date |
|---|---|---|---|---|---|
| WO 9824935 | A | 11-06-1998 | AU | 722819 B2 | 10-08-2000 |
| | | | AU | 5515198 A | 29-06-1998 |
| | | | EP | 0960214 A1 | 01-12-1999 |
| | | | WO | 9824935 A1 | 11-06-1998 |
| | | | US | 6190857 B1 | 20-02-2001 |
| US 4895809 | A | 23-01-1990 | NONE | | |
| US 6197599 | B1 | 06-03-2001 | NONE | | |
| WO 0210338 | A | 07-02-2002 | AU | 7908201 A | 13-02-2002 |
| | | | WO | 0210338 A2 | 07-02-2002 |